

CLAIMS

1. A system of parsing unstructured or partially structured data; said system processing at least portions of said data in an incremental manner.
- 5 2. The system of Claim 1 wherein said processing in an incremental manner comprises multiple parsing steps, each parsing step performed by consulting an inference engine.
- 10 3. A knowledge base for use in association with the system of Claim 1 ~~or Claim 2~~, said knowledge base analyzing said data at one or more predefined levels of analysis.
4. The knowledge base of Claim 3 wherein said levels include a level of analysis at a lexico-grammatical level.
- 15 5. The knowledge base of Claim 3 wherein said levels include a level of analysis at an orthographic level.
6. The knowledge base of Claim 3 wherein said levels include a level of analysis at a semantic level.
7. The knowledge base of Claim 3 wherein said levels include a level of analysis at a contextual level.
- 20

09883123 061501
T05T90" E2T8860

8. The knowledge base of Claim 3 wherein said knowledge base uses a knowledge representation language which embodies linguistic theory.
9. The knowledge base of Claim 8 wherein said linguistic theory is that of systematic functional linguistics.
10. The knowledge base of Claims 8 ~~or 9~~ wherein said linguistic theory enables the complete representation of all possible forms of said data.
11. The knowledge base of Claim 10 wherein said data is attribute data.
12. The knowledge base of Claim 11 wherein said attribute data is name and address data.
13. A method of parsing an attribute data set; said method comprising incrementally refining elements of said data set until a predefined level of meaning is determined.
14. The method of Claim 13 wherein said step of incrementally refining said elements includes execution of an elaboration operator.
15. The method of Claim 13 wherein said step of incrementally refining said elements includes execution of an encapsulation operator.

09883123-061501

16. The method of Claim 13 wherein said step of incrementally refining said elements includes execution of an enhancement operator.
17. The method of Claim 13 wherein said step of
5 incrementally refining said elements includes execution of an entailment operator.
18. The method of Claim 13 wherein said step of incrementally refining said elements includes execution of an extension operator.
- 10 19. The method of any one of Claims 13 ~~through to 18~~ wherein a best-first searching algorithm is utilized.
20. The method of any one of Claims 13 ~~to 18~~ wherein a look-ahead algorithm is utilized.
21. The system of any one of Claims 1 ~~to 18~~ wherein an
15 inference strategy is utilized.
22. A system for processing an unstructured or partially structured set of data so as to obtain a set of structured data; said system comprising a parser engine in communication with a knowledge database.

09883123.061501

23. The system of Claim 22 wherein said parser engine is
reliant on data in the form of knowledge retained in
said knowledge database.
24. The system of Claim 22 ~~or Claim 23~~ further including a
temporary data store associated with said parser engine.
25. The system of Claim 24 further including a data block
identifier which provides input to said parser engine.
26. The system of Claim 25 wherein said data block
identifier breaks said set of unstructured data into a
plurality of data blocks for input to said parser
engine.
27. The system of Claim 26 wherein said parser receives
consecutive ones of said data blocks and performs a
first association step on said data blocks based on
knowledge derived from said knowledge database so as to
derive a first postulated categorization of said data
blocks and storing said data blocks thereby categorized
in said temporary storage means.
28. The system of Claim 27 wherein said parser engine
performs a confirmation step on said data blocks stored
in said temporary storage means so as to either confirm
or reject its categorization of said data blocks.

FOI b7E b7C b7D b7F b7G b7H b7I b7J b7K b7L b7M b7N b7O b7P b7Q b7R b7S b7T b7U b7V b7W b7X b7Y b7Z

29. The system of any one of Claims 22 ~~through to 28~~ wherein said knowledge base includes knowledge about the information structures of identifying attribute objects.
30. The system of any one of Claims 22 ~~through to 29~~ wherein
5 said knowledge database includes knowledge about an association between patterns and the identifying attribute objects they represent.
31. The system of any one of Claims 22 ~~through to 30~~ wherein a precedence of alternative solutions has been
10 precompiled in said knowledge database thereby to allow best-first searching to be performed by said parser engine.
32. The system of any one of Claims 22 ~~through to 31~~ wherein said parser engine utilizes a best-first searching
15 algorithm.
33. The system of any one of Claims 22 ~~to 32~~ wherein said parser engine utilizes a look-ahead algorithm.
34. The system of any one of Claims 22 ~~to 33~~ wherein said parser engine utilizes an inference strategy.
- 20 35. The system of Claim 1 ~~or Claim 2 or any one of Claims 22 to 34~~ wherein said data comprises attribute data.

09883123.061501
TOST90.E28860

36. The system of Claim 35 wherein said attribute data comprises name and address data.

37. NEW DEPENDS ON 22
38. NEW " " 37

09883123.061501
TOST90" E2TE8860